

PCT

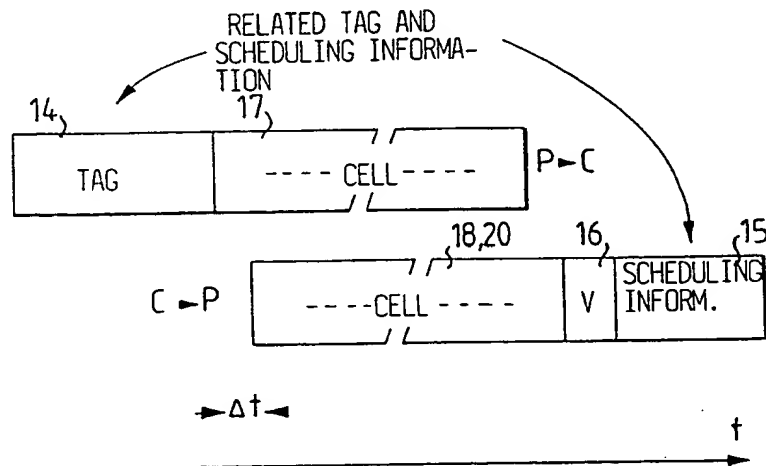
WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁵ : H04L 12/56	A1	(11) International Publication Number: WO 95/01031 (43) International Publication Date: 5 January 1995 (05.01.95)
(21) International Application Number: PCT/SE94/00599 (22) International Filing Date: 17 June 1994 (17.06.94) (30) Priority Data: 9302176-4 23 June 1993 (23.06.93) SE (71) Applicant: TELEFONAKTIEBOLAGET LM ERICSSON [SE/SE]; S-126 25 Stockholm (SE). (72) Inventor: SANDQUIST, Peter, Waldemar; Glasmästarvägen 9, S-125 31 Älvsjö (SE). (74) Agents: ROSENQUIST, Per, Olof et al.; Bergenstråhle & Lindvall AB, Box 17704, S-118 93 Stockholm (SE).	(81) Designated States: AU, BR, CA, CN, FI, JP, KR, NO, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report.</i>	

(54) Title: A CELL SWITCH AND A METHOD FOR ROUTING CELLS THERE THROUGH



(57) Abstract

A cell switch comprises a plurality of switch ports (2.a), each one connected to a number of user terminals, for sending and receiving cells of the user terminals, which cells contain payload and addressing information. A switch core (4) interconnects the switch ports for enabling communication between them. Tagging means are provided in the switch ports for creating, based upon the addressing information of a cell (17) received from a user terminal, a tag (14) belonging to the cell and containing routing information enabling routing of the cell through the switch core. Means are provided for sending tags to the switch core in advance of corresponding cells, one tag for each cell. Scheduling means in the switch core receive the tags sent in advance and create, based upon the routing information included therein, scheduling information regarding sending times of the respective cells from the switch ports for scheduling such that during these sending times the cells can be sent to the destination ports without conflict arising with other cells. The scheduling information is sent to the switch ports of origin of the cells, and the cells are retained in the switch ports until their respective scheduled sending times appear.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

A cell switch and a method for routing cells
therethrough.

5

Technical field of the invention

The present invention relates to a cell switch comprising a plurality of switch ports, each one connected to a number of user terminals, for sending and receiving cells of said user terminals, said cells containing payload and addressing information, a switch core, interconnecting said switch ports for enabling communication between them, tagging means in the switch ports for creating, based upon said addressing information of a cell received from a user terminal, a tag
10 belonging to the cell and containing routing information enabling routing of the cell through the switch core.

The invention furthermore relates to a method for routing cells through a cell switch which comprises a plurality of switch ports, each one connected to a number of user
20 terminals, for sending and receiving cells of said user terminals, said cells containing payload and addressing information, a switch core, interconnecting said switch ports for enabling communication between them.

Description of related art

25 In a switch of the kind indicated above cell buffers may be arranged so as to avoid loss of cells which simultaneously arrive to the switch and shall pass a resource common to the cells, e.g. an output. In known systems cell buffers need to be located in the switch core for attaining this object. In
30 order to attain a high efficiency a great number of cell buffers in the switch core are therefore required.

By locating cell buffers only in the switch ports the magnitude of the core and thereby its cost could be reduced. This would be advantageous in that a low cost for the switch
35 core reduces the total cost for scaleable systems which have been configured with only a few number of ports. The switch cost would thus be more modular.

One problem is, however, how to avoid conflict in the switch core so that the need of cell buffers in the core may be eliminated.

5 In EP 268,259 there is described a switch with a multiple stage packet transfer between switch inputs and outputs. A tag identifies packet destination for routing the packet to a suitable output while avoiding transmission conflicts. Buffering is provided for packets which may not immediately be transferred due to conflicts, an upper limit being set for
10 the stay of the packet in the packet transferring network.

The packets are handled in several steps and the tags are modified between the steps, but the tags are not sent in advance for scheduling of the sending times of the packets from the input ports.

15 In US 4,630,258 there is described a packet transfer node with a number of input and output ports. One single queue is connected to each input port for storing and transferring packets to all output ports. Conflicts may appear between packets directed to different ports. A multiple port memory
20 is connected between the input and output ports. A control logic is arranged to i.a. transfer packets from the memory to output ports identified in corresponding routing tags on the basis of a predetermined priority scheme. A buffer logic arbitrates tags arriving to the same input ports and
25 competing for the same output port. Furthermore there is carried through a random choice of packets directed to the same output port without need for the packet to wait during an undetermined time.

No tagging is carried through in the input ports.
30 Existing tags follow the transferred packets and are thus not sent in advance.

In US 4,621,359 there is described packet transfer where a packet transfer node handles packets containing routing tags indicative of output port destinations. The aim is to
35 provide a balance of the number of packets addressed for each of the output ports. New tags are therefore generated for redistributing the output port load. The new tags are

modified in order to obtain a predetermined output priority scheme.

There is no sending in advance of tags for providing a time controlled sending scheme for subsequent packets.

5 In US 4,623,996 there is described a packet transfer node which handles packets including routing tags indicative of output port destinations. A plurality of queue switches are individually connected between a plurality of input ports and a plurality of queue sets including a plurality of queues
10 which store and transfer applied packets as a function of the output port destination. The queues of each queue set are connected to different output arbitrators which control the routing to a particular output port.

Tags are not sent in advance for obtaining a time
15 controlled sending scheme for following packets.

Summary of the invention

One object of the invention is to solve, in a switch of the kind identified by way of introduction, the problem with conflict in the core in such a way that the need of cell
20 buffers in the core is eliminated.

The switch according to the invention comprises means for sending tags to the switch core in advance of corresponding cells, one tag for each cell, scheduling means in the switch core for receiving said tags sent in advance and, based upon
25 said routing information included therein, for creating scheduling information regarding sending times of the respective cells from the switch ports for scheduling such that during these sending times the cells can be sent to the destination ports without conflict arising with other cells,
30 means for sending said scheduling information to the switch ports of origin of the cells, delay means for retaining the cells in the switch ports until their respective scheduled sending times appear, and means for sending the cells into the switch core in their respective scheduled sending times.

35 The method according to the invention comprises the steps of creating, based upon said addressing information of received cells, a respective tag for each received cell

containing routing information enabling routing of the cell through the switch core to a desired destination port,

sending said tags to the switch core in advance of corresponding cells,

5 creating, based upon said routing information included in said tags sent in advance, scheduling information regarding sending times of the respective cells from the ports for scheduling such that during these sending times the cells can be sent to the destination ports without conflict arising
10 with other cells,

sending said scheduling information to the ports of origin of the cells,

retaining the cells in the ports until their respective scheduled sending times appear, and

15 sending the cells into the core in their respective scheduled sending times.

According to the invention conflict is thus avoided by the switch ports in advance providing routing information to the core. Thereby a suitable feedback to the ports may be
20 obtained so that cells may be sent in advance for avoiding simultaneous passing of common resources.

The cells are directed through the core by means of tags. A tag is a routing information preceding each cell. Each tag contains routing information which is not related to the
25 immediately arriving cell, but to some cell following thereafter, the cells being thus delayed in the ports. Processing of the routing information is carried through in the core, whereafter a suitable scheduling information is fed back to the ports.

30 The cells are transferred in adjacent fixed length intervals called transmission cycles and are sent through the core while using tags including internal cell destinations. The cell destinations may be provided, in a way known per se, as bits in a bitmap, whereby a cell can have several
35 destinations (group addressing). Tagging is carried through in the ports. Each cell is preceded by a tag. The tags are transferred during the cell transmission cycles from the

ports to the core. In the opposite direction, i.e. from the core to the ports, scheduling information is transferred and a flag for valid cell is transferred instead of the tags during the transmission cycles.

5 The phase location between incoming and outgoing transmission cycles is somewhat different so that a cell in a transmission cycle from a port to the core has time to pass through the core and without any extra delay may be located in a transmission cycle from the core to the port.

10 Cells are delayed in the ports a determined least number of transmission cycles after corresponding tags have been sent to the core so that time is reserved for processing for avoiding conflicts.

15 Conflict is avoided by distribution in time of the cells. The core schedules all cells and informs the ports so that the ports are forced to transmit the cells in certain transmission cycles. Cells competing for a common resource are transferred in different transmission cycles. A tag without routing information indicates to the core that no new
20 cell is waiting for being scheduled. The solution according to the invention results in a small core which may be used independently of the magnitude of the system within certain restrictions with a good cost efficiency. Furthermore systems being small from the beginning may easily be upgraded without
25 loss of the original investment.

Brief description of the drawings

The invention will now be described more closely below with reference to the attached drawings, on which

30 Fig. 1 illustrates the topology of an exemplary cell switch in which the invention is used,

Fig. 2 shows a tag with its cell routing bitmap,

Fig. 3 shows the format of the scheduling information,

Fig. 4 shows a cell consisting of a label and useful load,

35 Fig. 5 shows the phase shift between the transmission cycles for port-to-core and core-to-port, and the relation between the tag and scheduling information related to the

tag,

Fig. 6 shows the localization of a tag and its corresponding cell which is delayed by two transmission cycles,

5 Fig. 7 shows a block forming a part of a scheduler, and controlling the scheduling sequence and time priority,

Fig. 8 shows a state machine forming part of the scheduler and being included in a control block,

10 Fig. 9 shows a block forming part of the scheduler and creating lock pulses for memories for scheduling data,

Fig. 10 shows a memory for arriving tags,

Fig. 11 shows a data switch choosing between one of the rows from the memory for arriving tags and presenting it on the outputs,

15 Fig. 12 shows a circuit included in several places in the memory for scheduling data,

Fig. 13 shows an instance of the memory for scheduled data,

Fig. 14 shows an instance of a conflict control block,

20 Fig. 15 shows the scheduler including a number of the blocks shown in Fig.7-14,

Fig. 16 shows a delay means which detains the cells in the ports and send them to the core in the scheduled transmission cycles,

25 Fig. 17 shows a block included in several instances in the delay means,

Fig. 18 shows a state machine forming part of a block according to Fig. 17 of a delay means in the respective port,

30 Fig. 19 shows a scenario during which cells are connected through a switch of the suggested type with use of the invention.

Detailed description of preferred embodiments

Fig. 1 schematically shows a type of cell switch in which the invention can be used.

35 The switch includes a number of ports $2.n$ and a switch core 4 with a number of inputs $6.n$ and outputs $8.n$. For the sake of simplicity there are shown in Fig. 1 only four ports

2.1-2.4, four inputs 6.1-6.4, and four outputs 8.1-8.4. The ports 2.n form the outer interface of the switch. A plurality of user terminals may be connected to each port.

Through two oppositely directed transmission channels
5 10.n and 12.n, respectively, each port 2.n is connected to a corresponding input 6.n and output 8.n, respectively, of the switch core.

As an example the switch may be of the coordinate switch type and non-blocking, i.e. it shall be able to connect
10 arbitrary combinations of inputs and outputs simultaneously. The invention is, however, not limited to any particular inner structure of the switch.

Cell buffers are required for avoiding loss of cells due to conflict in the outputs, i.e. when several cells destined
15 for the same output port are sent to the switch with time overlap.

According to an important feature of the invention all cell buffers are localized to the ports. Since the core lacks cell buffers particular measures must be taken for avoiding
20 conflict of cells. This is performed, in a way to be described more closely below, by scheduling the arrival of the cells to the core and letting them in so as to avoid conflict at the outputs of the core.

All ports work cell-synchronously and the cells are
25 transferred in adjacent fixed length intervals called transmission cycles. The cells are stored in memory elements while awaiting their respective sending times. In the system given as an example it is assumed that the cells have some form of addressing information which is collected and
30 translated by looking in tables to internal routing information in the form of bit maps, where the positions of the bits represent different outputs. These maps obtained from the translation table are called tags, which is an adopted concept. In known systems the tags are sent together
35 with their corresponding cells, whereas in the system according to the present invention the tags and the corresponding cells are separated.

Fig. 2 schematically illustrates a tag 14 with a cell destination bitmap, including four fields 14.1-14.4, which in the example given quite simply are imagined as corresponding to each one of the four columns of the switch core. Each field may contain a bit which, e.g. only if it is set to one, indicates that the corresponding cell is destined to the column in question.

As will be described more closely below there are sent during each transmission cycle from each port 2.n a tag and a cell, which do not correspond to each other. The tag has the double function of providing routing information to the switch core 4 and asking for a transmission cycle for the corresponding cell. The core 4 processes the demand and assigns a transmission cycle indicating a cell transmission time for the port 2.n. The port awaits this point of time and sends the cell in the assigned transmission cycle. During the wait period other demands and cells may be handled.

When the tags in a transmission cycle have been received there is started the process of scheduling all the requested cells. All tags and corresponding scheduling information are saved until the cells have been sent. The scheduling information (the assigned transmission cycle), being the result of the scheduling process, is sent during the immediately following transmission cycle from the core 4 to the port 2.n in question.

Fig. 3 illustrates schematically the format of the scheduling information which is generally designated 15. The representation is binary. The three last of the four fields shown correspond to a bit b0, b1 and b2, respectively, corresponding to the first, second and third, respectively, bit of a three bit word. A zero (all bits set to zero) indicates that access to the core is admitted after a delay of a determined least number of transmission cycles. All other numbers indicate further wait cycles. 5 indicates e.g. that the cell shall be delayed further 5 transmission cycles. During this time cells directed to other outputs may be transmitted. The first of the four fields shown comprises a

flag 16 for a valid cell. The flag is activated when the cell sent from the core 4 to the ports 2.n is valid and not an idle cell.

Fig. 4 shows a cell 17 consisting of a label 18 and a useful load 20. The label is only interpreted and is possibly manipulated by a high level protocol in the switch ports. No internal routing is performed directly on the basis of the label.

Fig. 5 shows the phase shift Δt between the transmission cycles for the cell 17 with tag 14, which does not belong to this cell, in the direction port-to-core and, respectively, for the same cell 17 with a scheduling information 15 corresponding to the tag, in accordance with Fig. 3, and flag 16 for a valid cell in the direction core-to-port.

Fig. 6 shows the localizing of a tag 14 and its corresponding cell 17 when the cell is delayed by two transmission cycles. In the present embodiment this is a determined least delay.

When the cells arrive in accordance with the scheduling information 15 the saved tags 14 and the scheduling information are used for establishing the correspondence between the cells and the output port destinations.

The scheduler structurally consists of a number of blocks shown in Figures 7-14, which are interconnected in accordance with Fig. 15.

The scheduler on its inputs, designated 21 in Fig. 5, receives the tags arriving from the ports - one from each port. These tags comprise routing information and demands for transmission cycles.

The scheduler processes all such demands from the ports, one at a time, during a transmission cycle. As the demands are processed, one after the other, it becomes successively less probable during a transmission cycle that outgoing transmission cycles may be found, in which a conflict cannot appear. For this reason the order of processing of demands is changed for each transmission cycle for attaining justice between the ports.

The scheduler takes care of all tags corresponding to the coming third - sixth transmission cycles. All tags are stored in memory positions representing respective relative sending times. As each transmission cycle progresses all tags are
5 shifted into storage locations representing more adjacent scheduling times. The tags shifted out of the scheduler are shifted into FIFOs, not shown, which delay the tags two further transmission cycles comprising the outgoing first and second transmission cycles. The storage locations thus always
10 reflect the current state as time lapses, and thus the relative transmission times have been changed.

The input for "serial loading" of the storage location, that contains tags belonging to the most remote transmission cycle is fed with "not scheduled" (zeros). The output from
15 the said FIFOs is a map over the crossing points which shall be activated in each transmission cycle.

Each arriving tag is checked against earlier scheduled tags with respect to the risk for conflict. All processing of an arriving tag is performed simultaneously. Provided that at
20 least one outgoing transmission cycle has been found, in which conflict does not appear, the tag is stored in the storage location, that represents the earliest such transmission cycle.

After a complete sequence of processing of all tags of an
25 incoming transmission cycle some or all of the cells corresponding to these tags may have been scheduled. Their respective transmission times have been established during the processing and are supplied to a device that organizes the transmission of scheduling information (the cell
30 transmission times) to the ports.

Failure in scheduling of a cell may be indicated by e.g. stating a relative transmission time which is greater than the available number of storage locations, instead of a correct transmission time.

35 A control circuit acting as an input circuit of the scheduler, and generally indicated 22, is shown in Fig. 7 and 15. The control circuit 22 contains a state machine 24, which

is shown in more detail in Fig. 8 together with the appurtenant state table 26 and state graph 28.

5 The control circuit 22 in a way to be described more closely below controls the scheduling sequence and time priority.

Referring to the state table 26, the state machine 24 has four rest states S05, S15, S25 and S35, in which it is sensible for an arriving start signal on an input START. The start signal arrives once for each transmission cycle when arriving tags are available. The state machine 24 then performs counting operations (modulo-4), whereafter a state appears in which an output SH (for SHift) is active. Thereafter the next rest state appears. During the counting states an output SC (for SChedule) is active. Each of the
10
15 four counting sequences have different start values as appears from the state table.

A D-flipflop 30 receives the start signal intended for the state machine 24 and synchronizes it with a clock signal K before it is being fed to the input START of the state
20 machine 24. The clock signal is also supplied to an inverter 32, which creates an antiphase clock for the clock input CLK of the state machine 24. The clock signal also is supplied to one input of an AND gate 34, the other input of which is connected for receiving the start signal, so that the gate 34
25 on its output 35 supplies a synchronized locking signal for arriving tags.

The signal SH is decoded by an inverter 36 (one bit decoder). The decoded signal consists of the signal SH and the output from the inverter 36. A decoder 38 connected to
30 the state machine 24 in the way shown activates one of its outputs Y0-Y3 when the activating signal appears on its input ENable which is thus performed during a counting sequence.

A priority encoder 40 encodes the highest order input signal which is activated on one of its inputs P0-P3. An
35 output signal on its output ENable is set to zero when none of its inputs is activated. A decoder 42 connected to the encoder 40 in the way shown, is identical to the decoder 38.

The priority encoder 40 and decoder 42 together form a block that activates only one of the outputs Y0-Y3, that corresponds to the highest order input that is activated.

5 The function of the control circuit shown in Fig. 7 will appear from the following description of the scheduler and its function in connection with Figures 9-15.

10 The tags arriving to the scheduler on an input 43 are stored in a memory 44 in accordance to Fig. 10. Each row in this memory represents a tag for a particular port 2.n, and each column represents an address to one of the output ports 8.n. The tags are stored in a row of the memory in D- flipflops 45 and are locked by the locking signal, which in accordance with the above appears on the output 35 of the AND gate 34 in the control block 22. The outputs of the memory 44
15 are connected to inputs of a data switch which is described below with reference to Fig. 11.

The output signals of the decoders 38, 42 are fed to row conductors 46 and column conductors 48, respectively, of an addressing block 50 shown in Fig. 9, that creates locking
20 pulses for a memory for scheduling data, which is described more closely below with reference to Figures 12 and 13. To the row and column conductors 38 and 42, respectively, the respective inputs of AND-gates 52 are connected, and therefore a combination of signals - one activated signal for
25 each decoder 38 and 42, respectively, - thus activates the output of one of these AND-gates. The block 50 also contains OR-gates 54, in which one input is connected to the output of the corresponding AND-gate 52, and the other input is connected to the output SHift of the control block 22.

30 The outputs of the OR-gates 54 are all activated when SHift of the control block 22 is activated. Whenever either the SHift signal or the decoder signal to an AND-gate 52 is activated, the output of the OR-gate 54 connected to this particular AND-gate 52 will thus be activated. The outputs of
35 the OR-gates 54 are connected to an input of a respective AND-gate 56. The other input of each AND-gate 56 is connected for receipt of the clock signal K, and its output forms one

of the outputs 58 of the addressing block 50, and therefore signals appearing on the later are synchronized with the clock.

5 The addressing block 50 clocks either a storage location of a tag memory 60, to be described more closely below, for scheduled demands (Fig. 13), that represents a particular input port and an outgoing transmission cycle, or all storage locations.

10 A data switch 62 shown in Fig. 11 has, as has been mentioned above, inputs 64 connected to the outputs of the memory 44, and control inputs 66, 68, 70, 72 connected to the output of the decoder 38. The data switch 62 emits one of the tags from the memory 44 on one of its outputs 73 when the appurtenant control input 66-72 is activated.

15 Fig. 12 shows a block 74 with a flipflop 75 and an input switch consisting of AND-gates 76, 77 and an OR-gate 78. The control inputs 80 and 82 are connected to one input of a respective one of the gates 76, 77. When one of the control inputs 80, 82 is activated and the flipflop 75 is clocked,
20 the data on the data input corresponding to the activated control input is locked and presented on the output 84 of the flipflop 75.

A group of blocks 74 of the kind shown in Fig. 12 is included in the memory 60, shown in Fig. 13, for scheduled
25 demands. Each row of the group can receive one tag.

Fig. 14 shows an instance of the conflict control block. The conflict control block checks one tag at a time from the memory of arriving tags against a corresponding instance of the memory for scheduled tags.

30 The outputs from the data switch 62 shown in Fig. 11 are connected to inputs 85 of each of four conflict checking blocks 86 according to Fig. 14. The bits of a singled out tag are fed to one input of NAND-gates 88, 90, 92, 94 in each conflict checking block. Each conflict checking block 86 is
35 also connected to a corresponding one of the four tag memories 60 for scheduled demands according to Fig. 13. The outputs 96 for all rows and columns of the memories 60 are

rowwise connected to NOR-gates 98, 100, 102, 104, and
columnwise to OR-gates 106, 108, 110, 112. The NOR-gates 98-
104 set their outputs to zero when an arbitrary bit on their
respective rows is set to one. One of the rows is pointed to
5 by AND-gates 114, 116, 118, 120 and an OR-gate 122, the
output of which is connected to one of five inputs of and
AND-gate 124. The outputs from the OR-gates 106-112 are set
to one when an arbitrary bit on their respective columns is
set to one. The respective other input of the NAND-gates 88-
10 94 is connected to the output of a respective one of the OR-
gates 106, 108, 110 and 112, respectively. The NAND-gates 88-
94 therefore set their outputs to zero when their respective
tag bits received on the input 85 and an arbitrary one of the
bits in a column received from the corresponding OR-gate 106-
15 112 is set to one. The outputs of the NAND-gates 88-94 form a
respective one of the further inputs of the AND-gate 124. The
output of the AND-gate 124 is therefore set to one when there
are no bits set to one in the row pointed to, and tag bits
set to one on the input 85 do not exist on the same columns
20 as bits set to one in the tag memories. An arriving tag will
then possibly be able to be placed into the memory 60 for
scheduled demands, which appears from the following.

The outputs of all AND-gates 124 are connected to the
priority encoder 40 of the control block 22, and therefore
25 those of the inputs P0-P3 of the priority encoder will be set
to one, which correspond to transmission cycles, in which
cells can be successfully scheduled. The priority encoder 40
encodes the highest order input to a binary (2 bit) word and
an ENable signal. The output of the priority encoder 40 forms
30 the earliest outgoing transmission cycle (if there is one),
which does not cause cell conflict for the cell to be
scheduled. All outputs of the priority encoder 40 are
connected to the inputs of the encoder 42 and to flipflops,
not shown, and a device arranging for sending scheduling
35 information to the ports. The outputs of the encoder 42 are
connected to column choosing inputs of the addressing block
50, one output for each column. The columns of the addressing

block thus represent outgoing transmission cycles.

5 The arriving tag of a row which is processed, is fed to the inputs of all tag memories 60 for scheduled demands but from the addressing block 50 a locking pulse is arrived only to the storage location in the tag memories, that represents a port, from which the processed demand arrives, and an outgoing transmission cycle, in which no conflict will appear, whereby the tag is only clocked into this storage location.

10 At appearance of states in which SShift is activated, the decoder including the inverter 36 will bring the input selector included in the block in Fig. 12 to choose the other data input. These inputs and outputs 84 are so connected that a data shift is performed when the flipflops are clocked.

15 In each of the ports 2.n there is a delay means of the kind shown in Fig. 16. While awaiting scheduling information from the scheduler a cell arriving from outside the system is intermediately stored in a cell buffer, not shown, in the port. The cell is presented on the input "CELL IN" in all blocks 130 included in the delay means for further delay of cells, but is only stored in D-flipflops 152 in the block 130, the input "STORE" of which is set to one. Based upon scheduling information arriving from the core and presented for the inputs "SCHEDULING INFO", the cell on the output
20 "CELL OUT" is presented at the point of time, in which the cell shall be sent to the port, as will be described more closely below.

30 A block of the kind mentioned above is also shown in more detail in Fig. 17, a state machine 132 included therein being shown in Fig. 18 together with a corresponding state graph 130, a table 134 with state changing conditions and a state table 135.

35 Each of the blocks 130 has an output "FREE" which, when set to one, indicates that a corresponding buffer may store a new cell. These outputs from the blocks 130 are supplied to inputs P0, P1 and P2, respectively, of a priority encoder 136. The priority encoder 136 has an output EN (for ENable)

connected to an input of an AND-gate 138 which, in a way shown in Fig. 16, by another input is connected for receiving a flag "VALID SOURCE CELL FLAG". When set to one this flag indicates that a cell arriving from outside the system is not an idle cell. The output of the AND-gate 138 is connected to an input EN (for ENable) of a decoder 140. When the two inputs of the AND-gate are set to one, which occurs when the flag "VALID SOURCE CELL FLAG" and at least one of the inputs P0, P1 or P2 are set to one, the output on the AND-gate 138 is set to one. Thereby also some of the outputs Y0, Y1 or Y2 from the decoder 140 are set to one, corresponding to the highest order input which is activated by the inputs P0, P1 and P2 of the priority encoder 136. This has the result that a cell buffer 141 of that one of the blocks 130 that has its input "STORE" set to one, will be loaded with the cell, and that it guided by scheduling information received in the same block is presented on the output "CELL OUT" of the cell buffer in the transmission cycle, in which the cell shall be sent to the port. Through OR-gates 142 the cell is presented on the output 144 of the delay means, since only one buffer at a time can have outputs set to one.

The means included in each of the blocks 130 for presenting a cell from a cell buffer 141 in the right transmission cycle include the state machine 132, which provides a locking signal on the output "LATCH", counts down the number of transmission cycles that is indicated by the scheduling information, which is received on the inputs "SCHEDULING INFO", and provides a reading out signal on the output "DISPATCH" when count down has been performed.

The state graph 133 in Fig. 18 shows the states and the state transitions appearing when the state machine counts down time. The state S00 is a rest state in which the state machine is when the cell has been sent away and no new cell shall be stored in the cell buffer. The state S10 concerns reading out of the cell in the immediately following transmission cycle, whereas S20, S30 and S40 concern reading out after count down.

The table 134 with state changing conditions shows the rest of the state transitions. To the rest state S00 the state machine arrives from S00 (the same state), S10, S20, S30 or S40. To S10, S21, S32 and S43 the state machine
5 arrives either from the rest state or from the state S10 (the same state), S20, S30 or S40. To S10 the state machine arrives if the cell shall be sent in the immediately following transmission cycle, whereas arrival to S21, S32 and S43 occurs if the cell shall be delayed for further
10 transmission cycles.

In the state table 135 the status of the outputs are shown in different states. "FREE" is set to one only in the state S00. "LATCH" is set to one in the states S10, S21, S32 and S43. "DISPATCH" is set to one in the states S10, S20, S30
15 and S40.

For obtaining an antiphase clock for the state machine 132 the clock signal is sent to "CLOCK" via an inverter 146.

An AND-gate 148 and D-flipflop 150 are connected to the output "LATCH" and "DISPATCH", respectively, of the state
20 machine 132. The clock signal is also connected to the other input of the AND-gate 148 and the clock input of the D-flipflop 150, whereby synchronization occurs of the signals "LATCH" and "DISPATCH" with the clock. The output from the AND-gate 148 is connected to clock inputs of D-flipflops
25 152.1 - 152.n, which together form the cell buffer 141 and the data input D of which is connected to the input "CELL IN" of the block 130.

The output from the D-flipflop 150 is connected to an input of AND-gates 154.1 - 154.n, in which a second input is
30 connected to a respective output of the D-flipflops 152.1 - 152.n. The output of the AND-gates 154 is connected to the output "CELL OUT" of the block 130. Thus clocking of the cell buffer 129 is controlled via the AND-gate 148 by the signal "LATCH", and the presentation of the cell on the output "CELL
35 OUT" of the block 130 via the D-flipflop 150 by the signal "DISPATCH".

The clock pulses arrive synchronously and in phase with

outgoing transmission cycles. In the same transmission cycle as the scheduling information for a cell arrives, it and "STORE" must be stable at the point of time when the state machine is clocked.

5 In Fig. 19 there is shown an exemplary coupling scenario for demonstrating a possible behaviour of the system. Five complete consecutive transmission cycles for port-to-core and core-to-port are shown and indicated below in the Figure.

10 For indicating tag, cell and scheduling information in Fig. 19 the same concepts and reference characters are used as has been described with reference to and shown in Fig. 2-6. Different sets of interconnected tag, cell and scheduling information are distinguished on the one hand by addition of a different small letter, although common to the tag, cell and scheduling information, after the respective reference character, on the other hand by different line markings in the Figure.

20 Each tag is indicated by a sequence of four equally sized rectangular fields, corresponding to the fields 14.1-14.4 in Fig. 2 and, as therein, indicating a bit. The same is true for the scheduling information with fields corresponding to the fields 16, b2, b1, b0 in Fig. 3. Each cell is indicated, as in Fig. 5, with a lying and obliquely cut rectangle.

25 The rows A1 and A2 designate a transmission cycle for port-to-core and core-to-port, respectively, relating to a first port, which may be imagined to correspond to the port 2.1 in Fig. 1. The same is true for the rows B1 and B2, C1 and C2, D1 and D2 with respect to the ports 2.2, 2.3 and 2.4, respectively.

30 During the first transmission cycle there is supplied, as appears from line A1, a tag 14a to the core from the port 1. With the tag 14a also a cell 17 is supplied which, as has been said earlier, does not belong to the tag 14a. The tag 14a indicates a demand to send a cell to the port 3, by its third bit (14.3) being set to one, indicated by the line marking of the field in question. Since no other demands to send cells to the port 3 appear, the cell 17a belonging to

35

the tag 14a is scheduled a determined least delay as appears on line A2 by the scheduling information 15a from the core to port 1 in the transmission cycle 2 (all bits b2, b1 and b0 are set to zero). The least delay determined in the exemplary
5 system is two transmission cycles. Port 1 receives scheduling information 15a and accordingly sends the cell 17a to the core in the transmission cycle 3, as shown in line A1. The core indicates, line C2, in the transmission cycle 3 in the direction from the core to port 3 that the cell 17a is valid,
10 as appears from the flag 16a. In the same transmission cycle the core sends the cell 17a to the port 3 with a minimum dt of further delay and practically no buffering.

Three tags 14b, 14c and 14d in lines B1, C1 and C3, respectively, with conflicting cell destinations to port 1,
15 as appears from the respective first field (14.1 in Fig. 2), are transferred to the core during the transmission cycle 1. The tag 14c of the port 3 indicates that the cell belonging thereto has several destinations, viz. besides port 1 also port 4 (fields 14.1 and 14.4 in Fig. 2).

20 The conflict processing algorithm residing in the core has assigned different sending times to the cells from ports 2, 3 and 4 corresponding to the tags 14b, 14c and 14d with a common destination to port 1. Depending upon the state of the state machine 24 in the example, the tag 14b from the port 3
25 will be processed first in the core, cf. the scheduling information 15c in line C2, and thereby obtain the earliest sending time, and thereafter ports 4 and 2 follow. The cell 17c from the port 3 is thus sent to port 1 and 4 in the transmission cycle 3, cf. line C1 for direction to the core,
30 and lines A2 and D2 for direction towards the respective ports. The cell 17d together with a tag not belonging thereto is sent from the port 4 to the core, cf. line D1, and from there to port 1 in the transmission cycle 4, cf. line A2. The cell 14b together with a tag not belonging thereto is sent
35 from port 2 to the core, cf. line B1, and from there to the port 1 in the transmission cycle 5.

As the scheduled cell 17d of the port 4, with destination

port 1, waits for being sent further, a new cell 17e from port 4 with destination port 3 is scheduled. To the cell 17e is assigned the transmission cycle 5, cf. line D1 and C2, since earlier scheduled cell from port 4 shall be sent further in the transmission cycle 4. As the scheduled cell 17b in port 2, with destination port 1, is waiting for being sent further, there is scheduled and sent further another cell 17f from port 2 to port 3, cf. line B1 and C2.

Claims

1. A cell switch comprising
a plurality of switch ports (2.n), each one connected to
5 a number of user terminals, for sending and receiving cells
of said user terminals, said cells containing payload and
addressing information,
a switch core (4), interconnecting said switch ports for
enabling communication between them,
10 tagging means in the switch ports for creating, based
upon said addressing information of a cell (17) received from
a user terminal, a tag (14) belonging to the cell and
containing routing information enabling routing of the cell
through the switch core,
15 and further comprising
means for sending tags to the switch core in advance of
corresponding cells, one tag for each cell,
scheduling means in the switch core for receiving said
tags sent in advance and, based upon said routing information
20 included therein, for creating scheduling information
regarding sending times of the respective cells from the
switch ports for scheduling such that during these sending
times the cells can be sent to the destination ports without
conflict arising with other cells,
25 means for sending said scheduling information to the
switch ports of origin of the cells,
delay means for retaining the cells in the switch ports
until their respective scheduled sending times appear, and
means for sending the cells into the switch core in their
30 respective scheduled sending times.
2. A switch according to claim 1, wherein a tag and the
corresponding cell are transferred during different
transmission cycles from the corresponding port (2.n) to the
core (4).
- 35 3. A switch according to claim 2, wherein between the
transmission cycle for transmission of a tag and the
transmission cycle for transmission of a corresponding cell,

there is a time delay of a least number of transmission cycles, required for enabling processing for avoiding conflict.

5 4. A switch according to claim 2 or 3, wherein during each transmission cycle a scheduled cell is transferrable together with a tag of a cell not yet scheduled, in direction from the port to the core.

10 5. A switch according to any of claims 1-4, wherein a cell is sent together with the scheduling information of a later cell and a flag for valid cell from the core to the port.

6. A switch according to any of claims 1-5, wherein an invalid tag indicates to the core that no new cell waits for being scheduled.

15 7. A switch according to any of claims 1-6, wherein the tags sent from the ports and containing routing information each includes an implicit request for assignment of a transmission cycle.

20 8. A switch according to claim 7, wherein said scheduling means processes all such requests one at the time during a transmission cycle, the processing order being altered for each transmission cycle for attaining fairness between the ports.

25 9. A switch according to claim 7 or 8, wherein said scheduling means store all tags belonging to scheduled cells.

30 10. A switch according to any of claims 7-9, wherein all tags are stored in storage locations implicitly indicating the relative sending times for their corresponding cells, all tags being shifted for each new transmission cycle into new storage locations representing more near sending times so that the tags thus are always stored in storage locations reflecting the relative sending times.

35 11. A switch according to any of claims 7-10, wherein each tag arriving to the core is checked in parallel against tags of earlier scheduled cells regarding risk of conflict.

12. A switch according to any of claims 7-11, wherein failed scheduling of a cell is indicated by stating a

relative sending time for the cell which is greater than the available number of storage locations.

13. A method for routing cells through a cell switch which comprises

5 a plurality of switch ports, each one connected to a number of user terminals, for sending and receiving cells of said user terminals, said cells containing payload and addressing information,

10 a switch core, interconnecting said switch ports for enabling communication between them,

said method comprising the steps of

creating, based upon said addressing information of received cells, a respective tag for each received cell containing routing information enabling routing of the cell through the switch core to a desired destination port,

15 sending said tags to the switch core in advance of corresponding cells,

creating, based upon said routing information included in said tags sent in advance, scheduling information regarding sending times of the respective cells from the ports for scheduling such that during these sending times the cells can be sent to the destination ports without conflict arising with other cells,

25 sending said scheduling information to the ports of origin of the cells,

retaining the cells in the ports until their respective scheduled sending times appear, and

sending the cells into the core in their respective scheduled sending times.

30 14. A method according to claim 13, comprising the step of transferring a tag and the corresponding cell during different transmission cycles from the corresponding port to the core.

35 15. A method according to claim 14, comprising the step of providing, between the transmission cycle for transmission of a tag and the transmission cycle for transmission of a corresponding cell, a time delay of a least number of

transmission cycles, required for enabling processing for avoiding conflict.

16. A method according to claim 14 or 15, comprising the step of transferring during each transmission cycle a
5 scheduled cell together with a tag of a cell not yet scheduled, in direction from the port to the core.

17. A method according to any of claims 13-16, comprising the step of sending a cell together with the scheduling information of a later cell and a flag for valid cell from
10 the core to the port.

18. A method according to any of claims 13-17, comprising the step of indicating to the core by means of an invalid tag that no new cell waits for being scheduled.

19. A method according to any of claims 13-18, comprising
15 the step of providing each of the tags sent from the ports and containing routing information with an implicit request for assignment of a transmission cycle.

20. A method according to claim 19, comprising the step of processing all said requests, one at the time, during a
20 transmission cycle, and changing the order of such processing for each transmission cycle for attaining fairness between the ports.

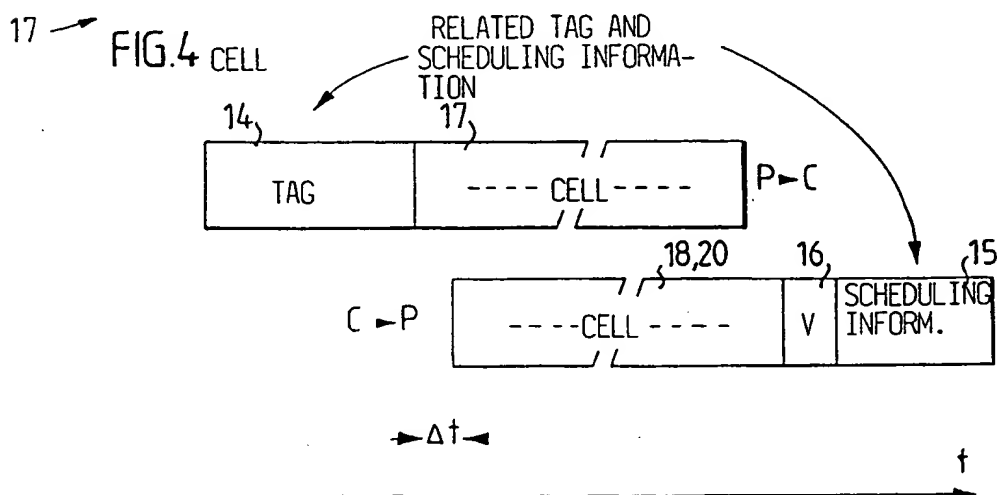
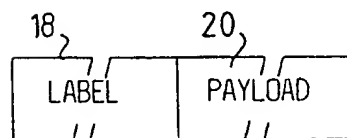
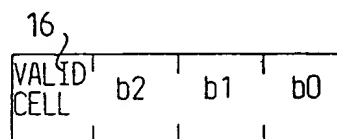
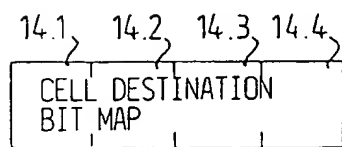
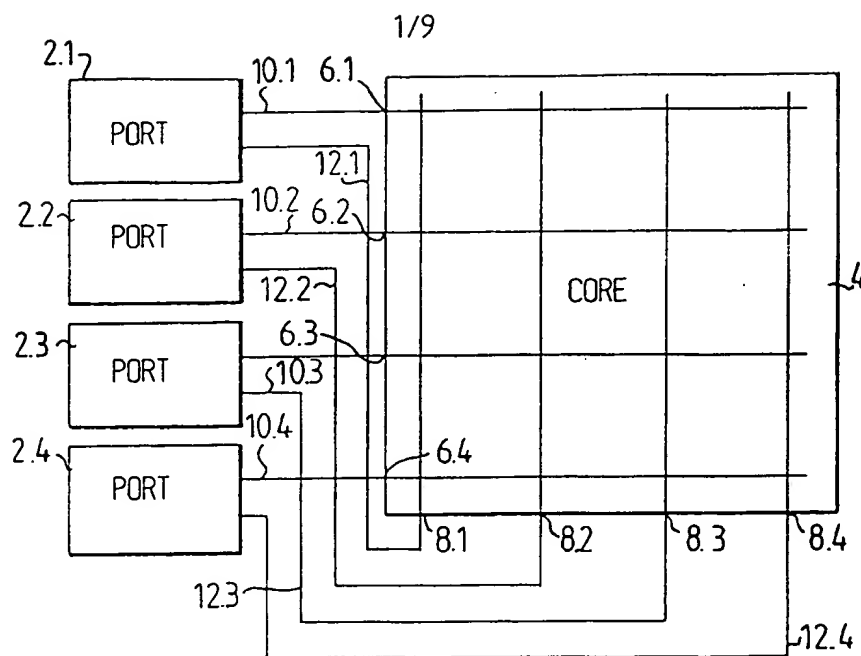
21. A method according to claim 19 or 20, comprising the step of storing all tags belonging to scheduled cells.

22. A method according to claim 21, comprising the steps
25 of storing all tags so that the storage locations implicitly indicate the relative sending times for their corresponding cells, and shifting all tags for each new transmission cycle into new storage locations representing more near sending
30 times so that the tags thus are always stored in storage locations reflecting the relative sending times.

23. A method according to any of claims 13-22, comprising the step of checking each tag arriving to the core in parallel against tags of earlier scheduled cells regarding
35 risk of conflict.

24. A method according to any of claims 13-23, comprising the step of indicating failed scheduling of a cell by

stating a relative sending time for the cell which is greater than the available number of storage locations.



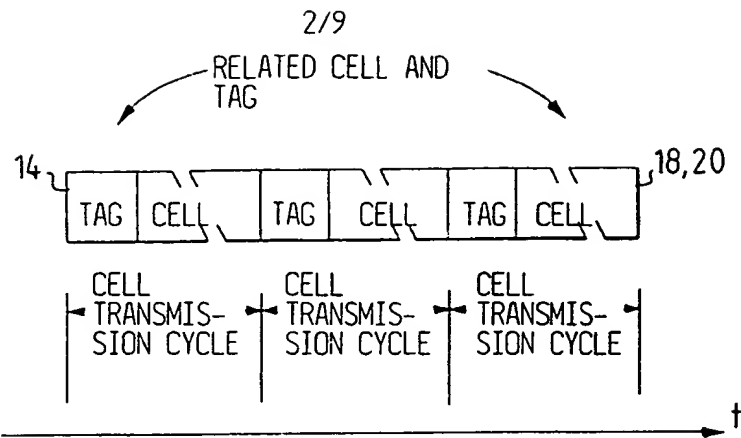


FIG. 6 RELATED CELL AND TAG WITH MINIMAL DELAY

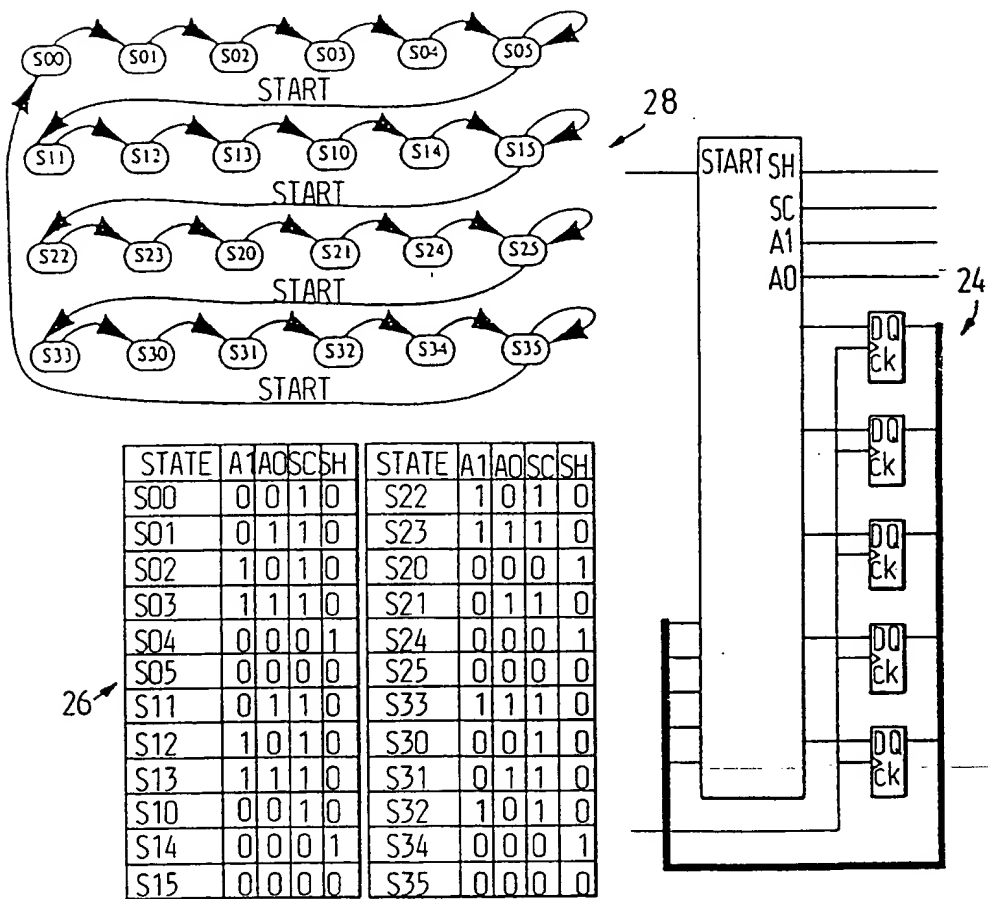
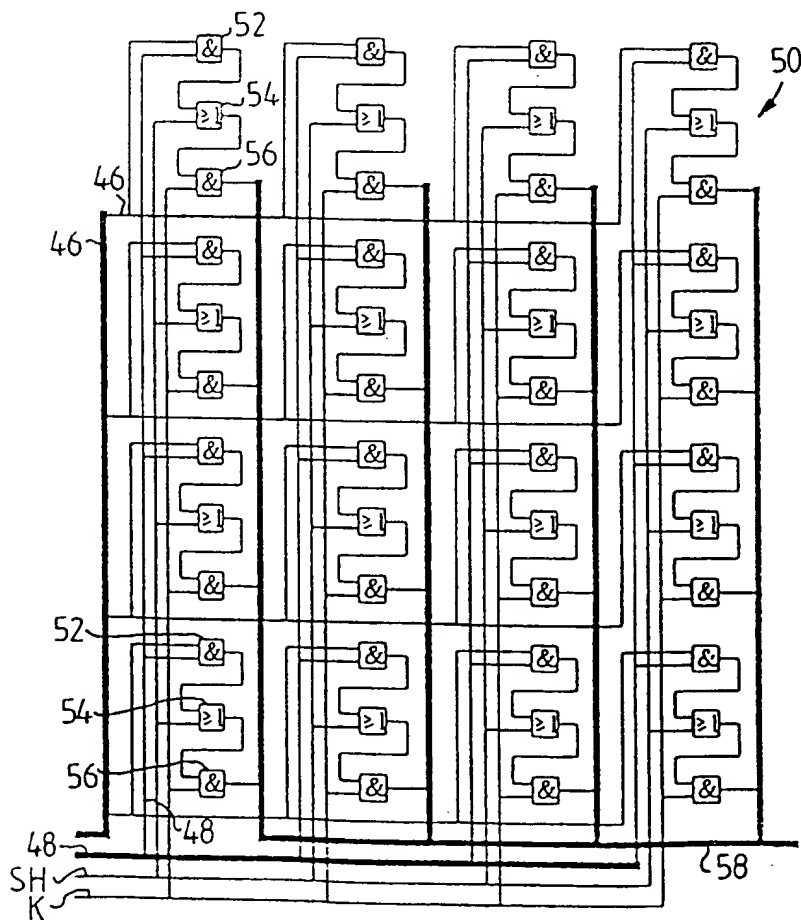
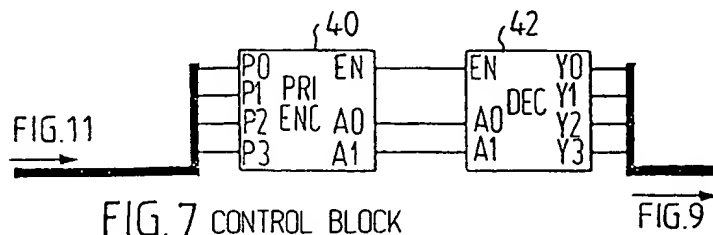
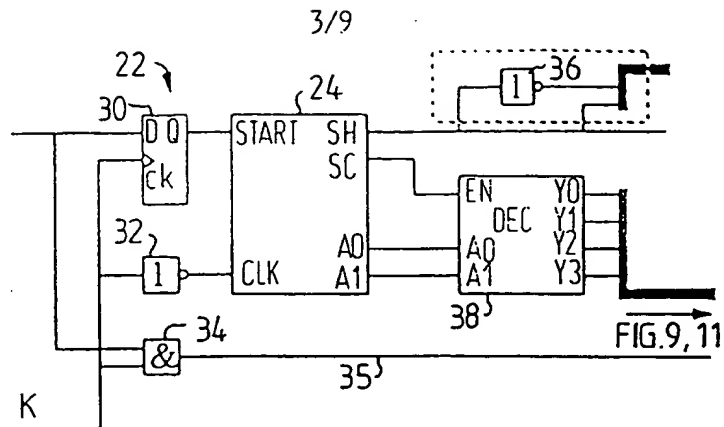
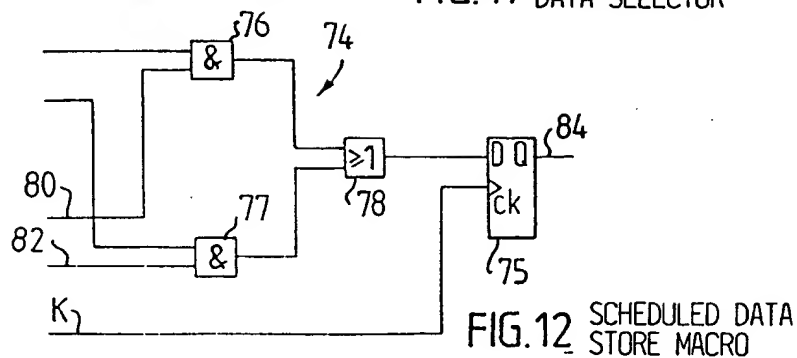
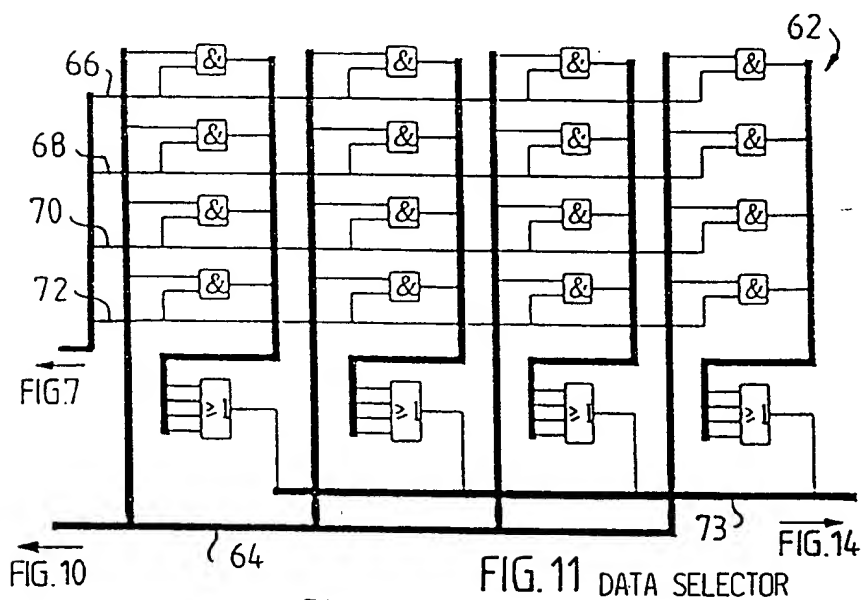
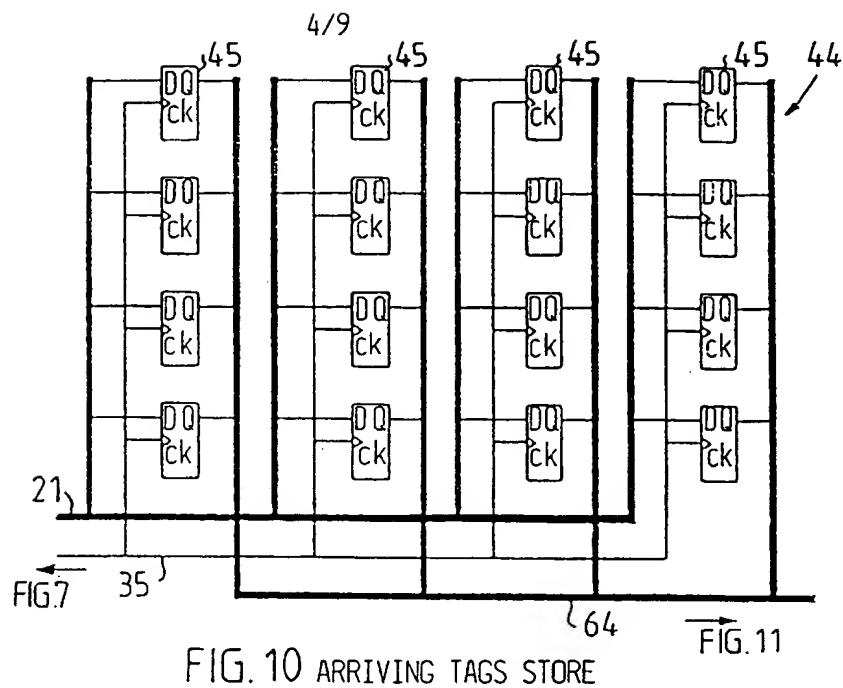


FIG. 8 STATE MACHINE

SUBSTITUTE SHEET



SUBSTITUTE SHEET



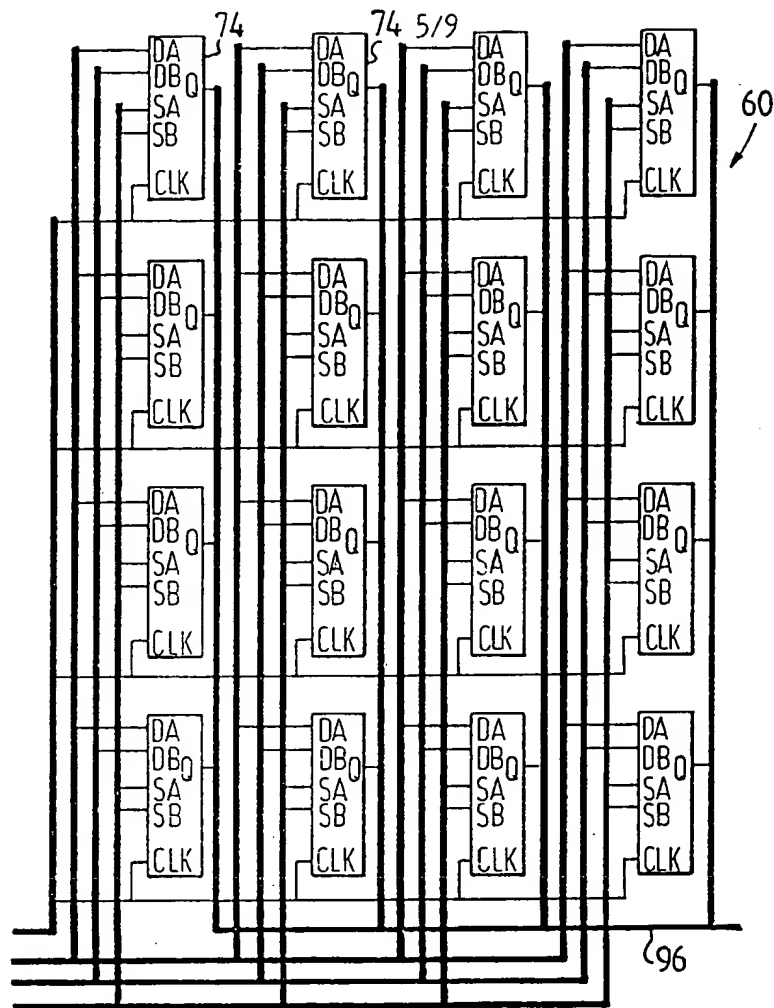


FIG. 13 SCHEDULED DATA STORE

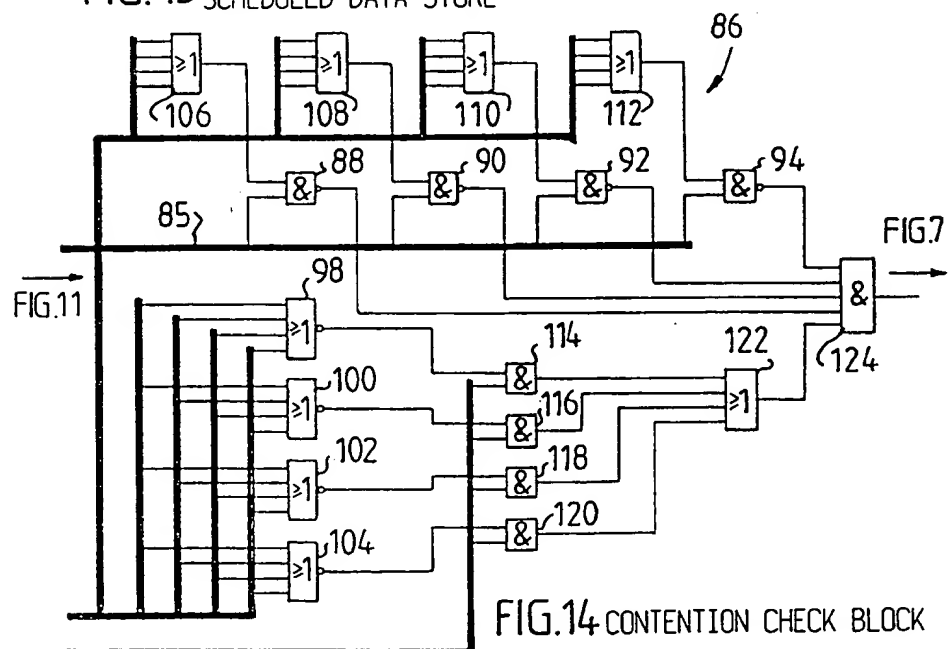


FIG. 14 CONTENTION CHECK BLOCK

6/9

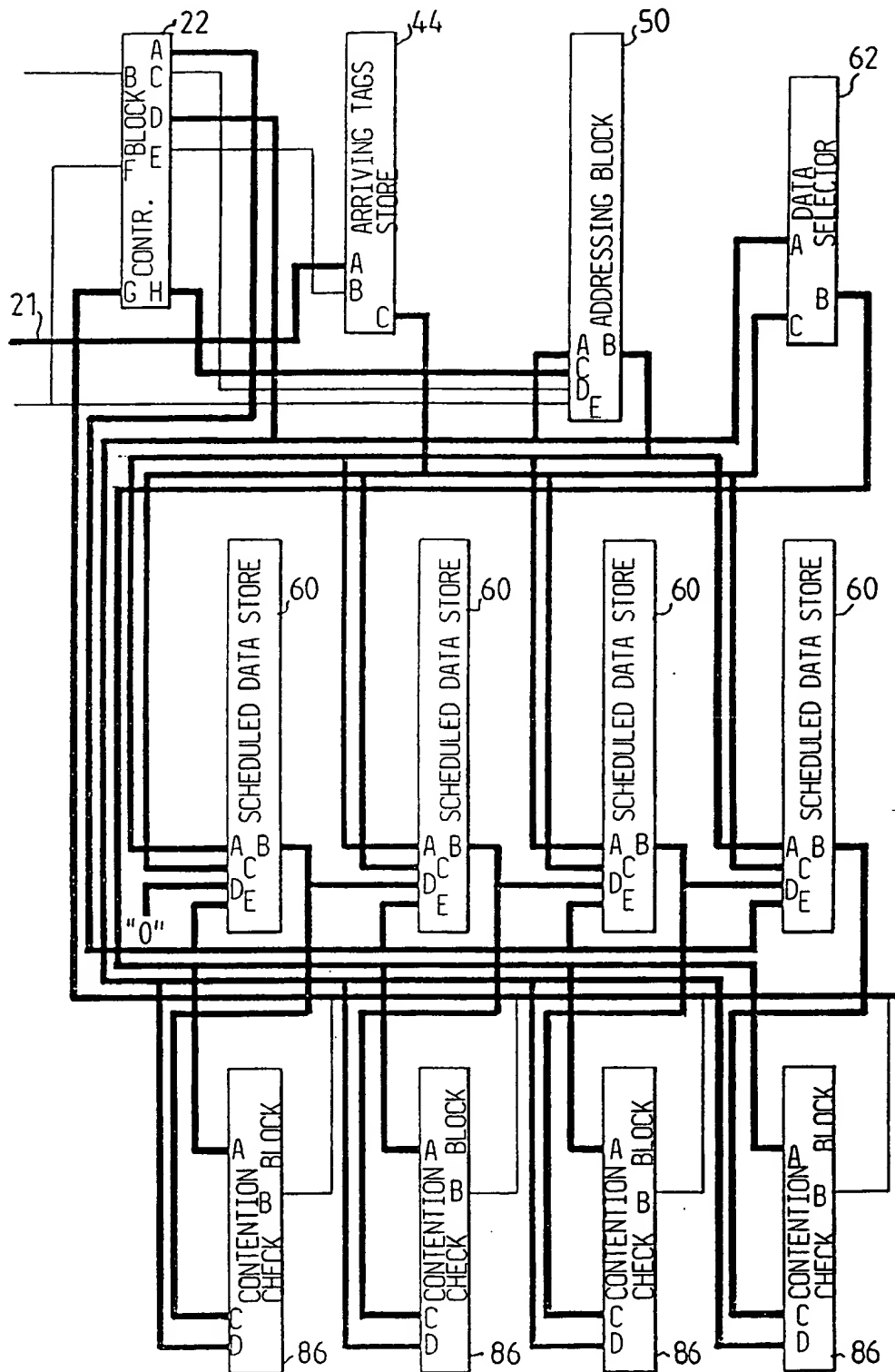


FIG.15 SCHEDULER

SUBSTITUTE SHEET

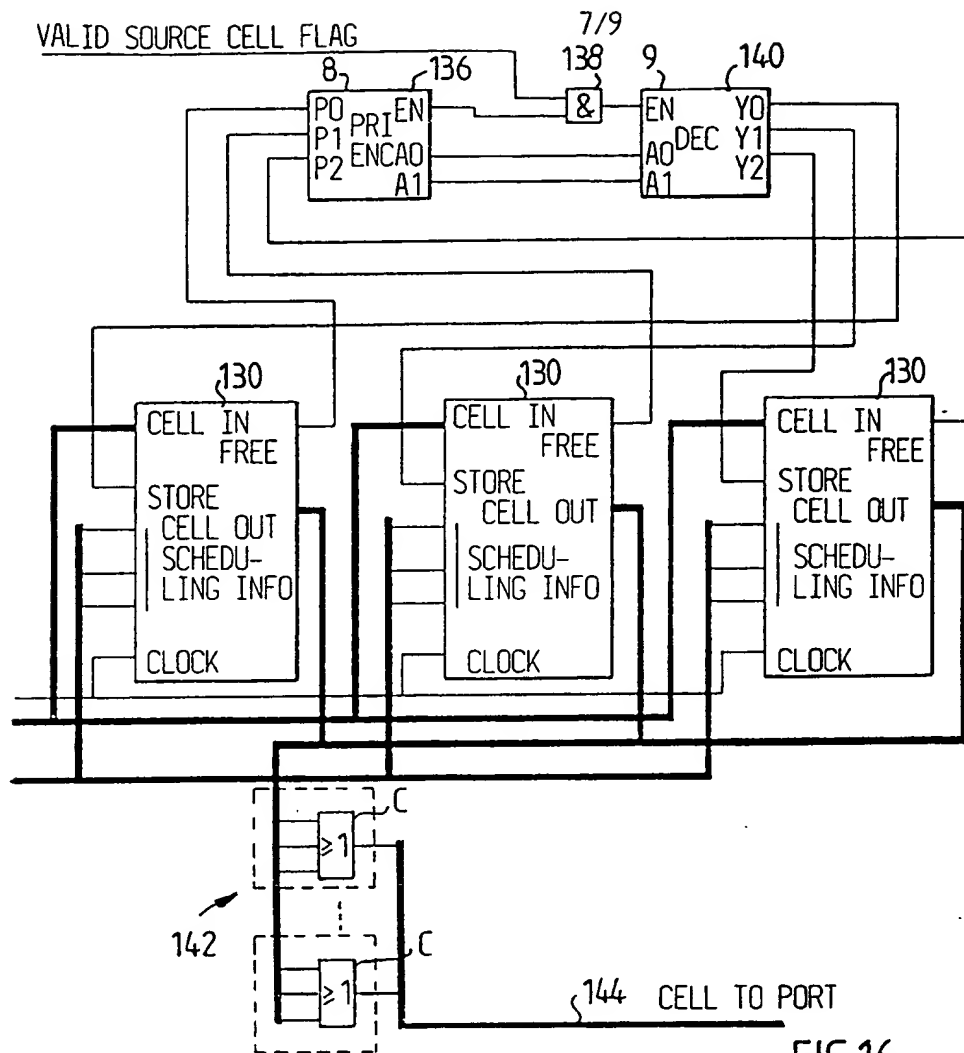


FIG.16

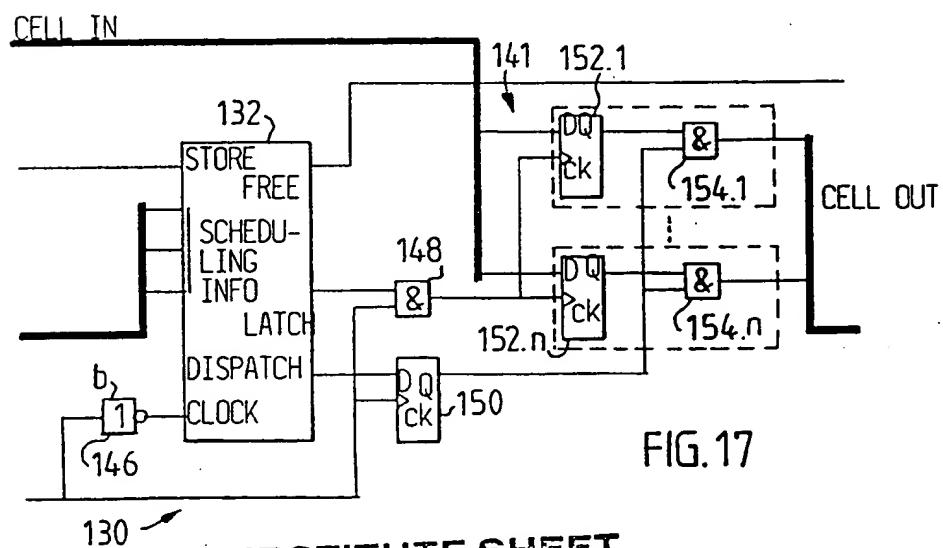
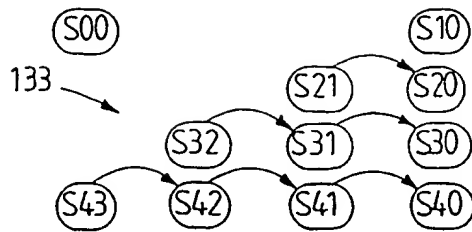


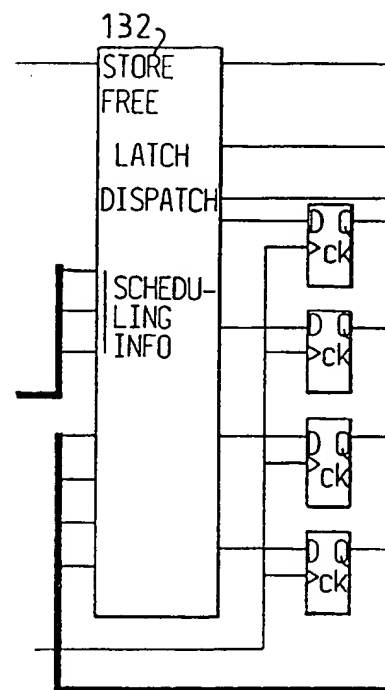
FIG. 17

8/9



Sxx>Sxx	STATE TRANSITION CONDITION
S00 S00	STORE
S00 S10	STORE & \bar{b}_2 & \bar{b}_1 & \bar{b}_0
S00 S21	STORE & \bar{b}_2 & \bar{b}_1 & \bar{b}_0
S00 S32	STORE & \bar{b}_2 & \bar{b}_1 & \bar{b}_0
S00 S43	STORE & \bar{b}_2 & \bar{b}_1 & \bar{b}_0
S10 S00	STORE
S10 S10	STORE & \bar{b}_2 & \bar{b}_1 & \bar{b}_0
S10 S21	STORE & \bar{b}_2 & \bar{b}_1 & \bar{b}_0
S10 S32	STORE & \bar{b}_2 & \bar{b}_1 & \bar{b}_0
S00 S43	STORE & \bar{b}_2 & \bar{b}_1 & \bar{b}_0
S20 S00	STORE
S20 S10	STORE & \bar{b}_2 & \bar{b}_1 & \bar{b}_0
S20 S21	STORE & \bar{b}_2 & \bar{b}_1 & \bar{b}_0
S20 S32	STORE & \bar{b}_2 & \bar{b}_1 & \bar{b}_0
S20 S43	STORE & \bar{b}_2 & \bar{b}_1 & \bar{b}_0
S30 S00	STORE
S30 S10	STORE & \bar{b}_2 & \bar{b}_1 & \bar{b}_0
S30 S21	STORE & \bar{b}_2 & \bar{b}_1 & \bar{b}_0
S30 S32	STORE & \bar{b}_2 & \bar{b}_1 & \bar{b}_0
S30 S42	STORE & \bar{b}_2 & \bar{b}_1 & \bar{b}_0
S40 S00	STORE
S40 S10	STORE & \bar{b}_2 & \bar{b}_1 & \bar{b}_0
S40 S21	STORE & \bar{b}_2 & \bar{b}_1 & \bar{b}_0
S40 S32	STORE & \bar{b}_2 & \bar{b}_1 & \bar{b}_0
S40 S42	STORE & \bar{b}_2 & \bar{b}_1 & \bar{b}_0

134



STATE	FREE	LATCH	DISPATCH
S00	1	0	0
S10	0	1	1
S21	0	1	0
S20	0	0	1
S32	0	1	0
S31	0	0	0
S30	0	0	1
S43	0	1	0
S42	0	0	0
S41	0	0	0
S40	0	0	1

135

FIG. 18

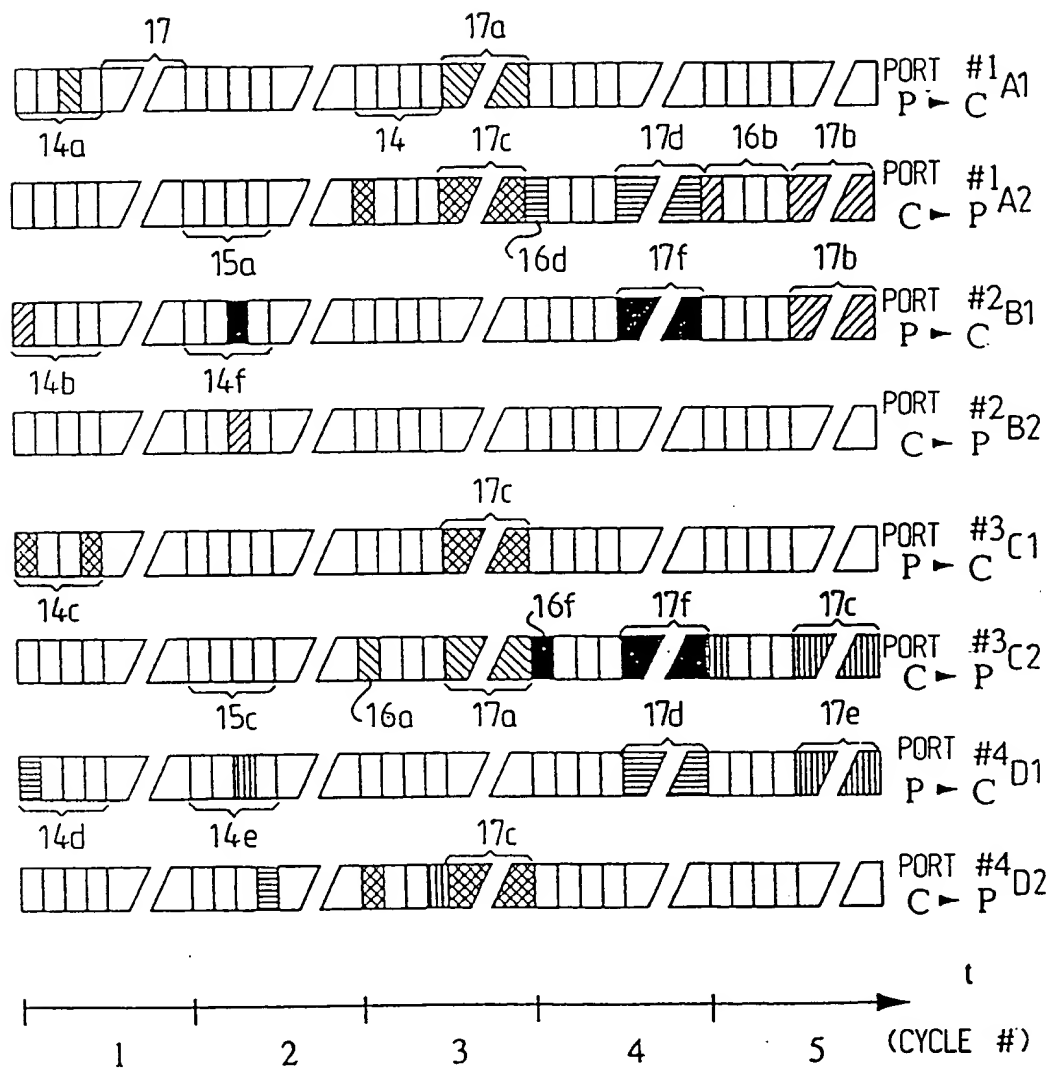


FIG.19

A. CLASSIFICATION OF SUBJECT MATTER		
5 IPC : H04L 12/56 According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols)		
5 IPC : H04L		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
SE,DK,FI,NO classes as above		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US, A, 4817084 (EDWARD ARTHURS ET AL), 28 March 1989 (28.03.89), column 2, line 42 - column 3, line 20 --	1-24
A	US, A, 5166926 (ARTURO CISNEROS ET AL), 24 November 1992 (24.11.92) --	1-24
A	US, A, 5038343 (GERALD LEBIZAY ET AL), 6 August 1991 (06.08.91) --	1-24
A	US, A, 4899334 (SHOGO SHIMIZU), 6 February 1990 (06.02.90) --	1-24
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
<p>* Special categories of cited documents:</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier document but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> <p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>"&" document member of the same patent family</p>		
Date of the actual completion of the international search		Date of mailing of the international search report
29 Sept 1994		10 -10- 1994
Name and mailing address of the ISA/ Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86		Authorized officer Göran Magnusson Telephone No. +46 8 782 25 00

INTERNATIONAL SEARCH REPORT

International application No.
PCT/SE 94/00599

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,A	US, A, 5255265 (KAI Y. ENG ET AL), 19 October 1993 (19.10.93) -- -----	1-24

INTERNATIONAL SEARCH REPORT
Information on patent family members

27/08/94

International application No.
PCT/SE 94/00599

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US-A- 4817084	28/03/89	CA-A- 1268843 DE-A- 3778911 EP-A,B- 0332611 JP-T- 1502380 WO-A- 8802961	08/05/90 11/06/92 20/09/89 17/08/89 21/04/88
US-A- 5166926	24/11/92	NONE	
US-A- 5038343	06/08/91	EP-A- 0405042 JP-A- 3038137	02/01/91 19/02/91
US-A- 4899334	06/02/90	EP-A- 0312917 JP-A- 1105641 JP-A- 1123550	26/04/89 24/04/89 16/05/89
US-A- 5255265	19/10/93	EP-A- 0569172 JP-A- 6037805	10/11/93 10/02/94